Genomic Sequence Analysis of the Mouse Desmoglein Cluster Reveals Evidence for Six Distinct Genes: Characterization of Mouse *DSG4*, *DSG5*, and *DSG6*

Neil Vincent Whittock

The desmosomal cadherins, comprising the desmogleins and desmocollins, are calcium-dependent transmembrane adhesion molecules that are essential for the cell adhesive role of desmosomes. Until recently, three mouse and three human desmoglein isoforms had been characterized that are expressed in a tissue- and differentiation-specific manner. Very recently, however, we revealed genetic evidence for the presence of a fourth human gene, *DSG4.* Here we present genetic evidence for the mouse *DSG4* homolog as well as two additional novel mouse desmoglein genes situated within the mouse cluster. We have named these two new genes *DSG5* and *DSG6*, both of which demonstrate close homology with mouse *DSG1.* Mouse *DSG4* comprises 16 exons spanning 36 kb of 18q, whereas *DSG5* and *DSG6* comprise 15 exons spanning approximately 33 and 37 kb, respectively, of 18q. The mouse desmoglein 4 transcript contains an open reading frame of 3123 bp, encoding a precursor of 1041 amino acids. The desmoglein 5 transcript contains an open reading frame of 3180 bp encoding a precursor of 1060 amino acid residues, and the desmoglein 6 transcript contains an open reading frame of 2733 bp encoding a precursor of 911 amino acid residues. Using mouse tissue cDNA we have demonstrated that mouse desmogleins 4, 5, and 6 are all expressed in the epidermis but are expressed during different times of mouse development.
Key words: Desmosome/cell-cell adhesion/desmosomal cadherins/pemphigus. J Invest Dermatol 120:970–980, 2003

The desmogleins and desmocollins, termed the desmosomal cadherins, are members of the cadherin superfamily. This superfamily comprises at least six subfamilies based on gene organization, protein domain composition, and phylogenetic analysis of their protein sequences (Nollet *et al*, 2000). The subfamilies comprise the type I classical cadherins, the type II atypical cadherins, desmocollins, desmogleins, protocadherins, and Flamingo cadherins. Initial experiments on bovine muzzle epidermis, a tissue rich in desmosomes, identified a constitutive transmembrane glycoprotein that was named ''desmoglein'' (also known as band 3 protein) (Schmelz *et al*, 1986). Using the antibodies raised against this bovine desmoglein further studies identified the cDNA and protein sequence of bovine desmoglein 1 (Koch *et al*, 1990; 1991). Similar to the classical cadherins, the desmosomal cadherins contain five homologous extracellular domains, a single transmembrane spanning domain, and a carboxy-terminal intracellular tail (Koch *et al*, 1990; Wheeler *et al*, 1991). The amino-terminal regions of the desmogleins and desmocollins make up the extracellular core domain of the desmosome (North *et al*, 1999). Calcium-dependent heterophilic adhesion occurs between the desmocollins and desmogleins of different cells, in addition to some homophilic adhesion (Chitaev and Troyanovsky, 1997; Marcozzi *et al*, 1998; Tselepis *et al*, 1998). In particular, studies have demonstrated that the first two extracellular domains (EC1 and EC2) of dcsmocol-lin are necessary for heterophilic binding to the first three extracellular domains of desmoglein (EC1—3) (Chitaev and Troyanovsky, 1997). Recent work on the adhesive ability of the first two extracellular domains of desmocollin 2 and desmoglein 2 has shown that desmocollin 2 can form homophilic interactions as well as calcium-dependent heterophilic reactions with desmoglein 2, whereas desmoglein 2 demonstrates a much weaker ability to form homophilic interactions (Syed *et al*, 2002). These results therefore suggest that both desmoglein and desmocollin isoforms are required in the same cell for optimal cell adhesion.

Within the desmosome the intracellular tails of the desmogleins and desmocollins bind the ''arm'' repeat domain of plakoglobin (Kapprell *et al*, 1988; Troyanovsky *et al*, 1993, 1994a; 1994b; 1996; Mathur *et al*, 1994; Ozawa *et al*, 1995; Wahl *et al*, 1996; Witcher *et al*, 1996; Andl and Stanley, 2001). Early studies suggested that, although desmocollins could bind only one molecule of plakoglobin, the desmogleins could bind six to seven plakoglobin molecules (Kowalczyk *et al*, 1996; Witcher *et al*, 1996). This is now in doubt as a recent study has shown that the desmogleins bind no more than one or two plakoglobin molecules in an isoform-spe-cific manner (Bannon *et al*, 2001).

Extensive work on several species including mice, cows, and humans revealed that there were three different isoforms of both the desmogleins and desmocollins although the desmocollins can exist in a and b forms (Buxton *et al*, 1993). Desmoglein 2 and desmocollin 2 appear to be the largest isoforms and are the essential desmosomal cadherins common to all desmosome-posscssing tissues, including simple epithelia, myocardium, and cell cultures (Koch *et al*, 1992; Theis *et al*, 1993; Schmidt *et al*, 1994). In contrast, the epidermal isoforms desmogleins 1 and 3, and desmocollins 1 and 3 are restricted to certain specialized epithelia (Koch *et al*, 1992; Schafer *et al*, 1994; 1996; Schmidt *et al*, 1994). Within the epidermis the desmosomal cadherin isoforms are expressed in a differentiation-specific manner, with desmoglein 1 and desmocollin 1 being upper spinous and granular layer, desmoglein 2 and desmocollin 2 being basal, and desmoglein 3 and desmocollin 3 being basal layer and first suprabasal layer (Arnemann *et al*, 1993; King *et al*, 1993; 1995; Theis *et al*, 1993; Schmidt *et al*, 1994; Nuber *et al*, 1995; Yue *et al*, 1995; Adams *et al*, 1998; Denning *et al*, 1998). The isoform of desmosomal cadherin present in individual des-mosomes relies upon whether the desmosome is in a region where the territories of the individual isoforms

Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, UK

Reprint requests to: Dr. Neil V Whittock, Molecular Genetics Laboratory, Royal Devon and Exeter Hospital, Old Pathology Building, Barrack Road, Exetet EX2 5DW, U.K.; Email: nwhittoc@hgmp.mrc.ac.uk

Received 13 December 2002; revised 25 January 2003; accepted 3 February 2003

overlap or not. In a study on desmocollins 1 and 3 it was shown that the ratio of proteins within an individual dcsmosome alters gradually from the basal to immediate suprabasal and then the upper suprabasal layer (North *et al*, 1996), demonstrating that epidermal cells are continually modified as they climb through the epidermis with a continual turnover of desmosomal cadherin protein. In mouse development the ratio of desmocollin 1 to desmocollin 3 expression throughout the different layers in the epidermis is crucial to establishing the correct pattern of differentiation, and that the establishment of the adult pattern of desmocollin expression corresponds to the adult pattern of epidermal stratification (King *et al*, 1996; Chidgey *et al*, 1997).

Recently, we have shown genetic evidence for an additional desmoglein isoform in the human genome, which we have termed desmoglein 4 (Whittock and Bower, 2003), which demonstrates highest homology to desmoglein 3. Here we reveal genetic evidence for the mouse desmoglein 4 homolog as well as two additional mouse desmoglein isoforms that we term desmoglein 5 and desmoglein 6 and we demonstrate their specific tissue expression, in addition to reviewing the properties of the mouse desmogleins.

MATERIALS AND METHODS

Reverse transcription polymerase chain reaction (RT-PCR) and cDNA expression studies

Total RNA was extracted from adult mouse skin using the Perfect RNA kit (Eppendorf AG, Cambridge, UK) as per the manufacturer's instructions. RNA was reverse transcribed using random hexamer primers with the Thcrmo-script cDNA kit (Invitrogen, Paisley, UK) as per the manufacturer's instructions. First strand cDNA from mouse skin and multiple mouse tissues (BD Clontech, Basingstoke, UK) was used for PCR using intron crossing primers. The gene-specific pnmers used for RT-PCR in cDNA expression studies of desmoglein 1 were DSG1F (forward-situated in exon 11) 5′-GGC ACT TCT TCC ACT GAG AA-3′ and DSG1R (reverse-situated in exon 12) 5′-CAG ATC AGC AGG AAT GGA ACC-3′. Primers used for amplification of desmoglein 2 were DSG2F (forward-situated in exon 12) 5′-GTG TGG CTG CAC AGT ATG AC-3′ and DSG2R (reverse-situated in exon 14) 5′-CAT GCC TTC CTT GAG CAT CG-3′. Primers used for amplification of desmoglein 3 were DSG3F (forward-situated in exon 12) 5′-CCT GAC AGT GTG TCA ATG TG-3′ and DSG3R (reverse-situated in exon 13) 5′-GGC TGA GCT CCT TCG ATT CC-3′. Primers used for desmoglein 4 expression studies were DSG4F (forward-situated in exon 4) 5′-GCG GGG ATT GAT CGG CCA CC-3′ and DSG4R (reverse-situated in exon 7) 5′-CTT GAT TCT GCA GTC ACA TTC-3′. Primers used for expression studies of desmoglein 5 were DSG5F (forward-situated in exon 11) 5′-CAA GGC ACT TCT ACT GTG GG-3′ and DSG5R (reverse-situated in exon 12) 5′-CAG ATC AGC AGG AAT GGA ACC-3′. Primers used for expression studies of desmoglein 6 were DSG6F (forward-situated in exon 9) 5′-CAG GTC AAG CTA CAA ACA AG-3′ and DSG6R (reverse-situated in exon 11) 5′-GTA CCA TGA TGA TTG TCC CTG-3′. Primers used to amplify glyccraldehyde-3-phosphate dehydrogenase (GAPDH) were as supplied in the mouse tissue kit (BD Clontech). PCR was performed in a 25 pi reaction mixture containing 10 pmol forward and reverse primers, 5 nmol each of dTTP, dCTP, and dATP, 3.8 nmol dGTP, 1.25 nmol dcaza GTP, 1.25 pi dimethylsulfoxide, 25 pmol betaine, 10 mM Tris–HC1, 50 mM KC1, 1.5 mM $MgCl_2$, and 1.0 U Amplitaq Gold Taq polymerase (Applied Biosystems, Warrington, UK). After an initial denaturation of 95°C for 12 min, 30 cycles were performed of 95°C for 1 min, 55°C for 1 min, and 72°C for 2 min, with a final extension step of 72°C for 10 min. The PCR products were examined by 1.5% agarose gel electrophoresis, purified using spin columns (Qiagen, Crawley, UK), directly sequenced using Big Dye terminators on an ABI 377 genetic analyzer (Applied Biosystems), and analyzed using Sequence Navigator 1.0.1 software (Applied Biosystems).

RESULTS

Sequence analysis identifies the mouse *DSG4* homolog

Analysis of the mouse genome sequence data *in silico* using the Blat algorithm (http://genome.ucsc.edu) to find homology for human *DSG4* (Whittock and Bower, 2003) was successful in characterizing the majority of the mouse desmoglein 4 gene. In addition, the position of the mouse homolog coincided with several predicted genes including chrl8-197 (Gcneid), SMUST00000037563 (Enscmbl), chr18.21.011.a (Twinscan), although some algorithms failed to predict a gene in this region (Genie and Fgencsh + +). The mouse promoter and exon 1 were determined by analysis of HumMus.18.14381 as described for the human gene (Whittock and Bower, 2003). Sequence analysis of the mouse desmoglein 4 cDNA indicates a transcript of at least 3.5 kb containing an open reading frame of 3123 bp. Genomic organization reveals that the mouse gene, like the human, is composed of 16 exons spanning approximately 36 kb of mouse chromosome 18. The exons varied from 36 bp (exon 2) to 768 bp coding (exon 16) in size, and the introns varied from 9900 bp (intron 1) to 152 bp (intron 13) in size (**Table I**). The mouse coding region shares 82% nucleotide identity with the human desmoglein 4 transcript (Whittock and Bower, 2003) and 55% with mouse desmoglein 1 (Mahoney *et al*, 2002), 51% with mouse desmoglein 2 (Mahoney *et al*, 2002), and 58% with mouse desmoglein 3 (Ishikawa *et al*, 2000) (**Table IV**). Similar to the human desmoglein genes, mouse *DSG4* is more closely related to mouse *DSG3* as they both consist of 16 exons, whereas mouse *DSG1* and *DSG2* comprise 15 exons. Sequence comparison for the human and mouse *DSG4* promoters reveals high homology and several common binding sites for sex determining region Y gene product (SRY), AP-1, CCAAT, enhancer binding protein, GATA binding factor, growth factor independence 1, HNF-3/Fkh homolog 3, nuclear factor 1, and homeodomain factor Nkx-2.5/csx.

Analysis of transcript, predicted protein, and tissue expression for mouse desmoglein 4

The 3123 bp open reading frame of the mouse desmoglein 4 cDNA encodes a precursor protein of 1041 residues with a predicted molecular weight of 114,449 Da and isoelectric point of 4.58 (**Fig 1**). Like the human desmoglein 4 protein this precursor comprises a 21 amino acid signal sequence followed by a 28 amino acid prosequence that includes an NXS/T glycosylation site and ends in a basic putative proteolysis sequence RRQKR. Cleavage at this sequence would result in a mature mouse desmoglein 4 unglycosylated peptide of 992 amino acids with a molecular mass of 108,598 Da and an isoelectric point of 4.55. The extracellular region of mouse desmoglein 4 (581 amino acids), by homology with other desmogleins, is divided into four domains EC1—4, followed by an extracellular anchor domain (EA). The conserved tripeptide sequence His-Ala-Val of the typical cadhcrins, which is thought to be involved in cell adhesion, is represented in mouse desmoglein 4 by the conservatively substituted tripeptide sequence Arg-Ala-Leu. Additionally conserved sequences in the extracellular domain of mouse desmoglein 4 and other mouse desmogleins include putative Ca^{2+} binding motifs (DXNDN and A/VXDXD) and two further potential N-glycosylation sites (NXS/T), in EC1 that is conserved in the equivalent positions of desmogleins 1–3, and in EA that is conserved only in desmoglein 3. The extracellular domain is followed by a 25 amino acid predicted transmembrane domain. The cytoplasmic domain of desmoglein 4 (386 amino acids) consists of an intracellular anchor domain, an intracellular cadherin-typical segment, an intracellular linker domain, a repeat unit domain, and a terminal domain. The repeat unit domain contains one NVXVTE repeat and a highly related NVXYAE predicting two repeats as found in desmoglein 3 (**Fig 1**). The overall identity at the protein level of mouse desmoglein 4 with its human homolog (Whittock and Bower, 2003) is 82% with 90% similarity. The identity of mouse desmoglein 4 with mouse desmoglein 1 (Mahoney *et al*, 2002) is 42%, with mouse desmoglein 2 (Mahoney *et al*, 2002) is 37%, and with mouse desmoglein 3 (Ishikawa *et al*, 2000) is 47% (**Table V**). The expression profile of the mouse desmoglein 4 transcript using RT-PCR demonstrates expression in adult epidermis and whole embryo from gestational days 7 to 17. No expression was detected, however, in skeletal muscle, brain, lung, liver, spleen, kidney, testis, or heart (**Table VI**).

Genomic sequence analysis reveals two further novel mouse Desmogleins

During analysis of the annotated mouse genome data on chromosome 18 (http://genome.ucsc.edu) for the mouse homolog of human desmoglein 4, several other predicted genes with homology to the

Table I. Exon–intron boundaries of the mouse desmoglein 4 gene *DSG4*

Exon	Exon size (bp)	5′ splice donor sequence	Intron size (bp)	3′ splice donor sequence
1	Noncoding+48	TTTATG**gt**aagt	9900	ttac**ag**GTGGTA
2	36	GTTGAG**gt**aatg	531	ctt**ag**GTAAAA
3	132	GCCAGA**gt**gagt	2259	ttgt**ag**ATTAGA
4	156	TTCCTG**gt**aagt	308	gccc**ag**ATCCAT
5	145	ACGCCA**gt**gagt	1924	ttat**ag**ATGCAC
6	167	AGAGAG**gt**aaag	1022	atta**ag**CAACAC
7	135	ACCTCA**gt**aagt	1583	tttc**ag**TACTCA
8	186	GTCAAG**gt**aagc	1957	tttc**ag**ATGCTG
9	272	TGTCAG**gt**atgt	1411	aaac**ag**GTATGT
10	140	TAGATG**gt**aaga	2610	tttc**ag**ATGGCT
11	219	TCAATG**gt**aagt	926	ctcc**ag**GAACCT
12	306	TGCTCT**gt**aagt	2133	ggce**ag**TGTCAC
13	140	GACAGG**gt**gagt	152	tttc**ag**GATGTA
14	64	CATCTG**gt**cagt	1306	ttcc**ag**AAATCT
15	209	TCCGAG**gt**aagc	4149	tctt**ag**AAAGCG
16	768+ noncoding			

desmogleins were detected. These included one gene between the mouse desmocollin gene cluster and mouse *DSG1*, with predicted genes chrlS-194 (Geneid), Dir18.21.007.a (Twinscan), and C18000206 (Fgenesh++), and another gene between *DSG1* and *DSG4*, with predicted genes chrl8-196 (Geneid), Chrl8.21.009.a (Twinscan), and C18000209 (Fgenesh ++). The genes predicted by the various algorithms demonstrated varying levels of homology with the known mouse desmoglein genes, *DSG1–4*. In addition, the predicted mRNAs and proteins showed a high level of homology to mouse desmogleins 1–4, specifically to desmoglein 1. To determine whether these predicted genes were expressed and to confirm the correct gene structure, overlapping RT-PCR was used on mouse control cDNA.

Analysis of transcript, predicted protein, and tissue expression for mouse desmoglein 5

Sequence analysis of the gene between *DSG1* and *DSG4*, that we name desmoglein 5, revealed a novel desmoglein with a transcript of at least 3.4 kb containing an open reading frame of 3180 bp. Exon-intron organization demonstrated that the new gene, *DSG5*, comprised 15 exons spanning approximately 33 kb of mouse chromosome 18. The exons varied from 36 bp (exon 2) to 1017 bp coding (exon 15) in size, and the introns varied from 291 bp (intron 7) to 11,159 bp (intron 1) in size (**Table II**). Nucleotide homology of the mouse desmoglein 5 open reading frame is 96% with mouse desmoglein 1 (Mahoney *et al*, 2002), 52% with mouse desmoglein 2 (Mahoney *et al*, 2002), 54% with mouse desmoglein 3 (Ishikawa *et al*, 2000), and 55% with mouse desmoglein 4 (**Table IV**). The 3180 bp open reading frame encodes a precursor protein of 1060 residues with a predicted molecular weight of 114,454 Da and isoelectric point of 4.72 (**Fig 1**). This precursor comprises a 21 amino acid signal sequence followed by a 28 amino acid prosequence that includes an NXS/T glycosylation site and ends in a basic putative proteolysis sequence RRQKR. Cleavage at this sequence would result in a mature mouse desmoglein 5 unglycosylated peptide of 1011 amino acids with a molecular mass of 108,628 Da and an isoelectric point of 4.59. Within the extracellular domain, the highly conserved tripeptide sequence His-Ala-Val of the typical cadherins is represented in mouse desmoglein 5 by the conservatively substituted sequence Arg-Ala-Leu. In addition to the conserved putative Ca^{2+} binding motifs (DXNDN and A/VXDXD), there are three further potential N-glycosylation sites (NXS/T) in the extracellular domain of mouse desmoglein 5, in EC1 that is conserved in the equivalent positions of all mouse desmogleins, in EC2 that is conserved in desmogleins 1 and 3, and in EA that is conserved only in desmogleins 3 and 4 (**Fig 1**). The extracellular domain is followed by a 25 amino acid predicted transmembrane domain and a cytoplasmic domain of which the repeat unit domain contains five NVXVT/RE repeats (**Fig 1**). The protein identity of mouse desmoglein 5 with mouse desmoglein 1 (Mahoney *et al*, 2002) is 94%, with mouse desmoglein 2 (Mahoney *et al*, 2002) is 34%, with mouse desmoglein 3 (Ishikawa *et al*, 2000) is 44%, and with mouse desmoglein 4 is 42% (**Table V**). The tissue expression profile determined by RT-PCR demonstrates that mouse desmoglein 5 is expressed in adult epidermis and in developing embryo day 17. Expression was not detected at embryonic days 7–15, or in heart, whole brain, spleen, lung, liver, skeletal muscle, kidney, or testis (**Table VI**).

Analysis of transcript, predicted protein, and tissue expression for mouse desmoglein 6

Sequence analysis of the gene between the desmocollin gene cluster and *DSG1*, which we name desmoglein 6, also revealed a novel desmoglein with a transcript of at least 3.0 kb containing an open reading frame of 2733 bp. Exon–intron organization demonstrated that *DSG6* also comprised 15 exons spanning approximately 37 kb of mouse chromosome 18. The exons varied from 36 bp (exon 2) to 723 bp coding (exon 15) in size, and the introns varied from 291 bp (intron 7) to 16,812 bp (intron 1) in size (**Table III**). Homology at the nucleotide level for the mouse desmoglein 6 open reading frame is 82% with mouse desmoglein 1 (Mahoney *et al*, 2002), 49% with mouse desmoglein 2 (Mahoney *et al*, 2002), 55% with mouse desmoglein 3 (Ishikawa *et al*, 2000), 55% with mouse desmoglein 4, and 81% with mouse desmoglein 5 (**Table IV**). The 2733 bp open reading frame of the mouse desmoglein 6 transcript encodes a precursor protein of 911 amino acid residues with a predicted molecular weight of 100,516 Da and isoelectric point of 4.84 (**Fig 1**). As for desmoglein 5 this precursor comprises a 21 amino acid signal sequence followed by a 28 amino acid prosequence that includes an NXS/T glycosylation site and ends in a basic putative proteolysis sequence RRQKR. Cleavage at this sequence would result in a mature mouse desmoglein 6 unglycosylated peptide of 862 amino acids with a molecular mass of 94,690 Da and an isoelectric point of 4.68. As well as the conserved RAL and putative Ca binding motifs (DXNDN and A/VXDXD), there are three further potential N-glycosylation sites (NXS/T), in EC1 that is conserved in all mouse desmogleins, in EC2 that is conserved in desmoglcins 1, 3, and 5, and in EC4 that is not conserved in any other mouse desmoglein (**Fig 1**). The cytoplasmic domain of mouse desmoglein 6 contains a repeat unit domain with four NV/ IXVT/RE repeats and a truncated terminal domain. The identity at the protein level of mouse desmoglein 6 with mouse desmoglein 1 (Mahoney *et al*, 2002) is 87%, with mouse desmoglein 2 (Mahoney *et al*, 2002) is 34%, with mouse desmoglein 3 (Ishikawa *et al*, 2000) is 42%, with mouse desmoglein 4 is 44%, and with mouse desmoglein 5 is 86% (**Table V**). The tissue expression profile for desmoglein 6 demonstrates expression in epidermis, liver, and testis, and in embryos at 17 d. Expression was not detected, however, in whole brain, skeletal muscle, lung, heart, spleen, and kidney, or at embryonic days 7–15 (**Table VI**).

Expression of mouse desmogleins 1–3

In addition to determining the tissue expression profiles of the new mouse genes, RT-PCR using specific primers was also used to determine

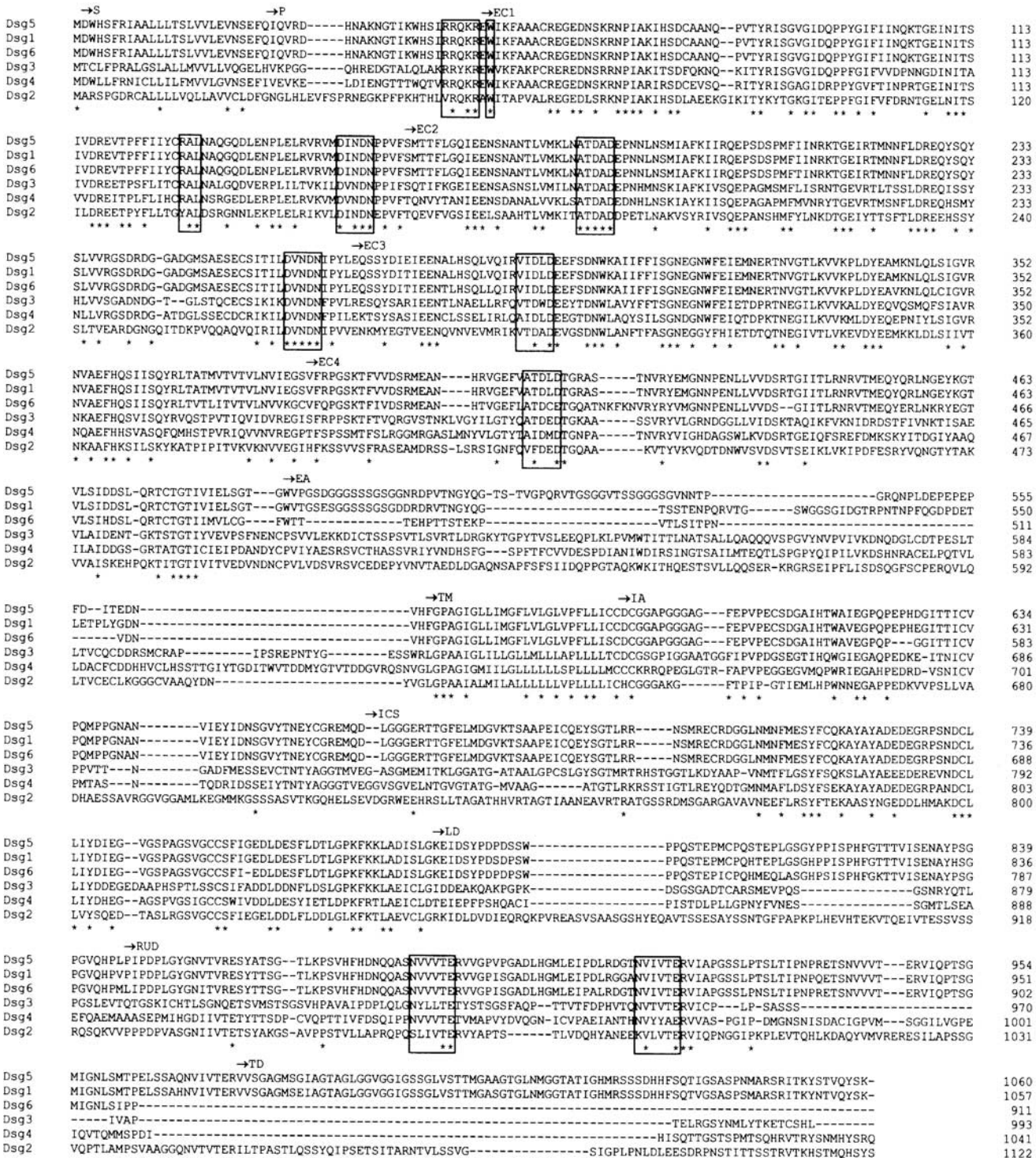

Figure 1. Multiple amino acid sequence alignment of mouse desmogleins 1–6. Horizontal arrows above the amino acid sequence show the beginning of each domain. Desmoglein 1 (Mahoney *et al*, 2002), desmoglein 2 (Mahoney *et al*, 2002), and desmoglein 3 (Ishikawa *et al*, 2000) are from published sequences. Conserved motifs for all mouse isoforms are boxed and conserved amino acids are represented by asterisks (*) under the sequence. S, signal sequence; p, preprotein domain; EC, extracellular domain; EA, extracellular anchor domain; TM, transmembrane domain; IA, intracellular anchor domain; ICS, intracellular cadherin-typical segment domain; LD, linker domain; RUD, repeat unit domain; TD, terminal domain.

the expression profiles of mouse desmogleins 1, 2, and 3. Results demonstrate that desmoglein 1 is expressed in epidermis and testis, and at embryonic days 15 and 17, but is not expressed in heart, whole brain, spleen, lung, liver, skeletal muscle, kidney, or at embryonic days 7 or 11 (**Table VI**). Results for desmoglein 2 demonstrate expression in all tissues tested (**Table VI**). Results for desmoglein 3 demonstrate expression in epidermis and at embryonic days 7–17, but not in heart, whole brain, spleen, lung, liver, skeletal muscle, kidney, or testis (**Table VI**).

Table II. Exon–intron boundaries of the mouse desmoglein 5 gene *DSG5*

Exon	Exon size (bp)	5′ splice donor sequence	Intron size (bp)	3′ splice donor sequence
1	Non-coding+48	TCCCTG**gt**aagt	11159	ttac**ag**GTGGTG
2	36	ATCCAG**gt**aaga	700	ttta**ag**GTAAGA
3	132	GCCAAA**gt**aagt	1144	Atac**ag**ATTCAT
4	156	TTCATT**gt**aagt	1591	gtct**ag**ATCTAC
5	145	ATGCAA**gt**aagt	2305	ttct**ag**ACACAC
6	167	AGAGAG**gt**aatc	1455	tttc**ag**CAATAT
7	135	TCATCT**gt**aagc	291	ttcc**ag**TATGAC
8	186	GTCAAG**gt**atgg	790	ttct**ag**CGCCTA
9	260	TGTTAG**gt**aaga	304	atgc**ag**ATATGA
10	140	TAGATG**gt**aaga	1574	ttac**ag**ATTCCC
11	339	TAGGAT**gt**aagt	5128	cttc**ag**TGGTTG
12	137	CACGAT**gt**aagt	310	ttttag**GGGATA**
13	73	ACTCAG**gt**aaga	2108	ttct**ag**GAGTTT
14	209	TGCCAG**gt**aagg	996	aatt**ag**AAAGCC
15	1017+noncoding			

Table III. Exon–intron boundaries of the mouse desmoglein 6 gene *DSG6*

Exon	Exon size (bp)	5′ splice donor sequence	Intron size (bp)	3′ splice donor sequence
1	Noncoding+48	TCCCTG**gt**aagt	16812	ttac**ag**GTGGTG
2	36	ATCGAG**gt**aaga	700	ttta**ag**GTAAGA
3	132	GCCAAA**gt**aagt	1097	atac**ag**ATTCAT
4	156	TTCATT**gt**aagt	1601	gtct**ag**ATGTAT
5	145	ATGCAA**gt**gagt	2346	ttct**ag**ACACAC
6	167	AGAGAG**gt**aatc	1456	tttc**ag**CAATAT
7	132	TGATCT**gt**aagc	291	ttcc**ag**TATGAC
8	186	GTCAAG**gt**atgg	2067	ttct**ag**GCGCTA
9	269	TGTTAG**gt**aaga	297	ttttag**ATACAG**
10	140	TACATG**gt**aaga	1581	ttac**ag**ATTCCC
11	186	TAGGAT**gt**aagt	2136	cttc**ag**TGGTTC
12	128	CCTGGT**gt**aagt	311	tttc**ag**GGGATA
13	73	ACTGAG**gt**aaga	2124	ttct**ag**GAGTTT
14	209	TGCGAG**gt**aagg	998	aatt**ag**AAAGCC
15	723 + noncoding			

Table IV. Nucleotide sequence homology of mouse desmogleins 1–6

	Dsg1	Dsg2	Dsg3	Dsg4	Dsg5	Dsg6
Dsgl	100	52	54	55	96	82
Dsg2	52	100	50	51	52	49
Dsg3	54	50	100	58	54	55
Dsg4	55	51	58	100	55	55
Dsg5	96	52	54	55	100	81
Dsg6	82	49	55	55	81	100

Mouse desmoglein 1, Mahoney *et al* (2002); mouse desmoglein 2, Mahoney *et al* (2002); mouse desmoglein 3, Ishikawa *et al* (2000). Values are expressed as percentage (%).

Table V. Amino acid sequence homology of mouse desmogleins 1–6

	Dsgl	Dsg2	Dsg3	Dsg4	Dsg5	Dsg6
Dsg1	100 (100)	34 (50)	43 (57)	42 (57)	94 (95)	87 (89)
Dsg2	34 (50)	100 (100)	39 (55)	37 (53)	34 (50)	34 (49)
Dsg3	43 (57)	39 (55)	100 (100)	47 (64)	44 (57)	42 (54)
Dsg4	42 (57)	37 (53)	47 (64)	100 (100)	42 (56)	44 (58)
Dsg5	94 (95)	34 (50)	44 (57)	42 (56)	100 (100)	86 (88)
Dsg6	87 (89)	34 (49)	42 (54)	44 (58)	86 (88)	100 (100)

Mouse desmoglein 1, Mahoney *et al* (2002); mouse desmoglein 2, Mahoney *et al* (2002); mouse desmoglein 3, Ishikawa *et al* (2000). Values are expressed as percentage (%). Values in parentheses are percentage similarity.

Desmoglein gene organization demonstrates conservation between man and mouse

The elucidation of the gene organization for the six mouse desmoglein genes is useful in determining conservation with the human genes. Mouse desmoglein 1 contains 15 exons (Mahoney *et al*, 2002) in agreement with that determined for the human gene (Hunt *et al*, 2001), although some reports have incorrectly determined that human desmoglein 1 comprises only 14 exons (Rickman *et al*, 1999; Frank *et al*, 2001). Both mouse and human desmoglein 2 comprise 15 exons (Whittock, unpublished data), although a recent report incorrectly demonstrated that the mouse gene comprised only 14 exons (Mahoney *et al*, 2002) whereby exons 12 and 13 were combined to make exon 12. The mouse and human desmoglein 3 genes (Frank *et al*, 2001) contain 16 exons, although some reports have stated that both contain only 15 exons (Silos *et al*, 1996; Koch *et al*, 1997; Ishikawa *et al*, 2000). Studies have also now demonstrated that both human (Whittock and Bower, 2003) and mouse (**Table I**) desmoglein 4 genes comprise 16 exons. Exon size conservation is also significant between the different genes. All of the genes comprise an exon 2 of 36 bp, exon 5 of 145 bp, exon 6 of 167 bp, and exon 8 of 186 bp (**Table VII**). In addition, exon sizes are conserved between desmogleins 3 and 4 for exons 1–6, 8–10, and 14. We also demonstrate here that both mouse desmoglein 5 and mouse desmoglein 6 comprise 15 exons (**Tables II, III**) although no human homologs exist. Desmogleins 1, 5, and 6 share exon sizes for exons 1–6, 8, 10, and 13–14. Therefore, it can be deduced from these observations that there is significant conservation of gene organization between different isoforms and between human and mouse. It can be concluded that the mouse desmoglein genes are a result through evolution of at least two stages of gene duplication yielding four genes with 15 exons and two genes with 16 exons.

Comparison of human and mouse desmoglein cluster arrangement demonstrates evolutionary deletion of two mouse genes

The human desmoglein genes are clustered within a 250 kb (**Fig 2**) region in the following order *5′-DSG1-DSG4-DSG3-DSG2-3′* (Simrak *et al*, 1995; Whittock and Bower, 2003), whereas the mouse genes are clustered within a 400 kb region in the order *5′-DSG6-DSGl-DSG5-DSG4-DSG3-DSG2-3′* (**Fig 2**). The question therefore arises whether mouse genes *DSG6* and *DSG5* were deleted during evolution through two separate events or more plausibly whether mouse *DSG5* is in fact

the human homolog of *DSG1*. Nucleotide sequence homology for mouse desmoglein 1 (Mahoney *et al*, 2002) and mouse desmoglein 5 with human desmoglein 1 (Nilles *et al*, 1991) is 79% for both (**Table VIII**), and at the amino acid level is 78% for both with 85% similarity (**Table IX**). Of note, homology of mouse desmoglein 6 with human desmoglein 1 is 70 and 77% at the nucleotide and protein levels, respectively (**Tables VIII, IX**), demonstrating that mouse desmoglein 6 is not the homolog of human desmoglein 1. In addition, comparison of the mouse and human promoters reveals high homology for mouse desmogleins 1 and 5 with human desmoglein 1 (not shown). Therefore, because of this very high homology of mouse desmoglein 1 and 5 with human desmoglein 1 at the protein, transcript, gene, and promoter level it is very difficult to make a conclusion. In addition, the transcript expression profiles of mouse desmoglein 1 and mouse desmoglein 5 are not too dissimilar (**Table VI**). Of the tissues screened, mouse desmoglein 1 is expressed in epidermis and testis and at embryonic days 15 and 17, whereas mouse desmoglein 5 is expressed in epidermis and at mouse embryonic day 17. Arguments regarding the immunolocalization of mouse desmoglein 1 must now be questioned as the epitopes recognized by the antibodies used most probably cross-react with not only desmoglein 5 but mouse desmoglein 6 as well. New antibodies are required that exploit the minor amino acid sequences between these new desmogleins in order to correctly determine the cellular localizations of these proteins. In conclusion, based on the gene clusters of man and mouse it is more likely that mouse desmoglein 5 is the true genetic homolog of human desmoglein 1, and that the two centromeric genes, now known as *DSG6* and *DSG1*, were deleted from the desmocollin/dcsmoglein interlocus region.

Table VI. Isoforms of desmoglein in different mouse tissues as detected by RT-PCR

Tissue	Dsg1	Dsg2	Dsg3	Dsg4	Dsg5	Dsg6
Epidermis	+	+	+	+	+	+
Heart	–	+	–	–	–	–
Whole brain	–	+	–	–	–	–
Spleen	–	+	–	–	–	–
Lung	–	+	–	–	–	–
Liver	–	+	–	–	–	+
Skeletal muscle	–	+	–	–	–	–
Kidney	–	+	–	–	–	–
Testis	+	+	–	–	–	+
7 d embryo	–	+	+	+	–	–
11 d embryo	–	+	+	+	–	–
15 d embryo	+	+	+	+	–	–
17 d embryo	+	+	+	+	+	+

Comparison of mouse desmoglein isofbrms at the protein level demonstrate significant homology

Protein sequence comparison (**Figs 1, 3**) demonstrates that all mouse desmogleins contain a 21 residue signal sequence to direct them to the endoplasmic reticulum. The proscquence comprises 28 residues for isoforms 1 and 3–6, and 32 residues for isoform 2, all of which end in a five-residue cleavage recognition site with a consensus (R/V)R(Q/Y)KR. All of the mature proteins contain a tryptophan residue at position 2 and a potential N-glycosylation site (NXS/T) in EC1. The conserved tripeptide sequence in desmogleins 1 and 3–6 comprises Arg-Ala-Lcu, whereas desmoglein 2 comprises Tyr-Ala-Leu. Putative Ca^{2+} binding motifs (DXNDN and A/VXDXD) are conserved in all desmogleins at the end of ECl, the beginning and end of EC2, the beginning of EC3, and the beginning of EC4, although some conservative substitutions have occurred within the latter motif for desmoglein 6 (ATDCE). The EA domain demonstrates the least homology between the different mouse isoforms and is therefore ideal for exploitation in raising isoform-specific antibodies. Each isoform contains a 25 residue transmembrane domain followed by conserved intracellular anchor and intracellular cadherin segments. The repeat unit domains comprise repeats of approximately 30 residues containing the consensus sequence (N)(V/I/Y)(X)(V/L/Y)(T/A/R)(E). Desmogleins 1 and 5 contain five repeats, desmoglein 6 contains four repeats, and

Table VII. Comparison of the deduced exon sizes for the six mouse desmoglein genes

Exon	*DSG1*	*DSG2*	*DSG3*	*DSG4*	*DSG5*	*DSG6*
1	Non-coding+48	Non-coding+60	Non-coding+48	Non-coding+48	Non-coding+48	Noncoding+48
2	36	36	36	36	36	36
3	132	135	132	132	132	132
4	156	162	156	156	156	156
5	145	145	145	145	145	145
6	167	167	167	167	167	167
7	135	138	129	135	135	132
8	186	186	186	186	186	186
9	260	266	272	272	260	269
10	140	143	140	140	140	140
11	330	228	228	219	339	186
12	137	228	255	306	137	128
13	73	122	143	140	73	73
14	209	330	64	64	209	209
15	1017+coding	1020+noncoding	221	209	1017+noncoding	723+noncoding
16			657+noncoding	768+noncoding		

DSG1, Mahoney *et al* (2002); *DSG2* and *DSG3,* Whittock (unpublished data).

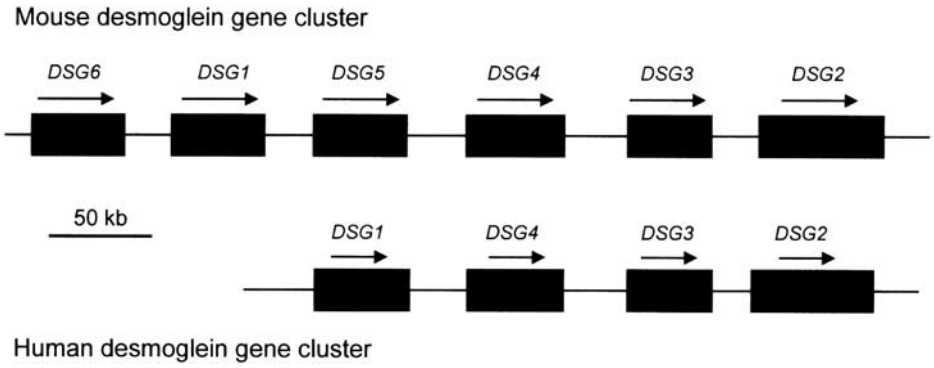

Figure 2. Schematic representation of the mouse and human desmoglein clusters. Mouse *DSG6* lies centromeric to *DSG1*, with *DSG5* and *DSC4* between *DSG1* and *DSG3*, with the entire mouse desmoglein complex occupying approximately 400 kb. The human desmoglein cluster occupies approximately 250 kb of 18q12.

Table VIII. Nucleotide sequence homology of mouse desmogleins 1, 5, and 6 with human desmoglein 1

	Mouse dsg1	Mouse dsg5	Mouse dsg6	Human dsg1
Mouse dsg1	100	96	82	79
Mouse dsg5	96	100	81	79
Mouse dsg6	82	81	100	70
Human dsg1	79	79	70	100

Mouse desmoglein 1, Mahoney *et al* (2002); human desmoglein 1, Nilles *et al* (1991).
Values are expressed as percentage (%).

Table IX. Amino acid sequence homology of mouse desmogleins 1, 5, and 6 with human desmoglein 1

	Mouse dsg1	Mouse dsg5	Mouse dsg6	Human dsg1
Mouse dsg1	100 (100)	94 (95)	87 (89)	78 (85)
Mouse dsg5	94 (95)	100 (100)	86 (88)	78 (85)
Mouse dsg6	87 (89)	86 (88)	100 (100)	77 (83)
Human dsg1	78 (85)	78 (85)	77 (83)	100 (100)

Mouse desmoglein 1, Mahoney *et al* (2002); human desmoglein 1, Nilles *et al* (1991).
Values are expressed as percentage (%).
Values in parentheses are percentage similarity.

desmogleins 3 and 4 contain two repeats. Human desmoglein 2 has been stated to contain six repeats (Schmidt *et al*, 1994) although the mouse homolog contains only two repeats that fit the consensus. Several very close motifs exist within mouse desmoglein 2, however (SLIVTE and KVLVTE), predicting that it may contain four repeats. Comparison of the overall sizes of the preproteins and mature proteins by residue content and molecular weight demonstrates that mouse desmoglein 2 is the largest of the isoforms (**Table X**) whereas desmoglein 6 is the smallest. In addition calculation of the isoelectric points for the isoforms demonstrates that desmoglein 2 has the most basic content whereas desmoglein 4 has the most acidic content, although all mature proteins have an isoelectric point in the range 4.55–5.01 (**Table X**).

GenBank accession number

Mouse desmoglein 4, 5, and 6 cDNA sequences have been deposited in the GenBank sequence database under Accession nos. AY191584, AY192158, and AY192159, respectively.

DISCUSSION

Here, we have demonstrated the presence of three new transcribed desmoglein genes, namely the mouse homolog of human dcsmoglein 4 and two novel mouse dcsmoglein isoforms that we term desmoglein 5 and desmoglein 6, and we show that they lie in a cluster of six genes. The identification of at least one additional desmoglein gene is not surprising as evidence has been obtained for at least one gene with homology to desmoglein 1 in the past (Puttagunta *et al*, 1994). Whilst isolating the bovine dcsmoglein 1 gene, evidence was found of polymorphism within the extracellular anchor domain. It is therefore likely that, as well as isolating bovine desmoglein 1, a bovine desmoglein 5 homolog had also been identified, although the authors could not conclude whether this region of polymorphism was allelic or represented a different gene (Puttagunta *et al*, 1994). It is unlikely that only bovine and now mouse species have additional desmoglein genes and it can therefore be concluded from these studies that a large number of species have at least one close homolog of desmoglein 1. The question then arises why some species have more genes and therefore more desmoglein proteins within their skin than in humans. Is this due to the fact that the living environment tor these animals is harsher and their skin is required to be tougher? Indeed, studies may show that species from an evolutionary earlier period have more than six desmoglein genes. Further studies on a wide range of species will be required to answer these questions.

The tissue expression of the mouse desmoglein isoforms during embryonic development may cast some light on their specific roles during gestation. Whereas mouse isoforms 2, 3, and 4 are expressed as early as 7 d gestation, isoform 1 is not expressed until gestational day 15, and isoforms 5 and 6 are not expressed until day 17 (**Table VI**). Stratification of the mouse body epidermis commences at around embryonic day 13.5 when it beconres mul-tilayered, and by embryonic day 14.5 the ventral epidermis is clearly stratified (King *et al*, 1996). Previous studies have shown that the expressions of desmoglein 1 as well as desmocollin 1 are closely related to keratinization of the epidermis (King *et al*, 1996). We have shown that transcripts for all six desmoglein isoforms are present in mouse epidermis and it therefore appears that the expression of mouse desmoglein isoforms 5 and 6 may well be closely related to keratinization of the epidermis and may indicate-that desmogleins 5 and 6 are expressed in the superficial layers along with desmoglein 1. The expressions of isoforms 2, 3, and 4, however, are linked with a much earlier process during mouse development, most probably the development of the basal cell layer of the epidermis, which may indicate that desmoglein 4 protein is expressed within the basal and immediate suprabasal cell layers along with desmoglein 3.

The involvement of the desmoglein isoforms has been implicated in human inherited and autoimmune disease as well as microbial

attack, all of which lead to loss of cell–cell adhesion. Striate palmoplantar keratoderma (MIM 148700) (Siemens, 1929; El Sayed and Bazcx, 1993) is an autosomal dominant genoderma-tosis characterized clinically by the development of linear and focal hyperkeratosis of the palms and soles during the first or second decade (Griffiths *et al*, 1998; Armstrong *et al*, 1999). The majority of cases of striate palmoplantar keratoderma are due to mutations within the extracellular domain of desmoglein 1 that lead to haploinsufficiency of the protein (Rickman *et al*, 1999; Hunt *et al*, 2001; Kljuic *et al*, 2003). Pemphigus foliaccus, fogo sel-vagem (endemic pemphigus foliaceus), and pemphigus vulgaris are autoimmune disorders characterized by blisters and erosions of the skin (Lever, 1965; Diaz *et al*, 1989). In pemphigus foliaceus and fogo selvagem the blisters occur beneath the stratum cor-ncum, whereas in pemphigus vulgaris the blistering is intracpi-dcrmal, occurring above the basal layer and with additional involvement of the mucous membranes. The autoantigens recognized by the circulating antibodies and T cells in pemphigus foliaceus and fogo selvagem have been identified as dcsmoglein 1, whereas those in pemphigus vulgaris recognize desmoglein 3 (Stanley *et al*, 1984; Labib *et al*, 1990; Amagai *et al*, 1991; Oursler *et al*, 1992; Olague-Alcala *et al*, 1994; Emery *et al*, 1995; Lin *et al*, 1997a; 1997b, 2000). Studies have shown that the autoantibodies in some patients with pemphigus foliaceus (7%) and pemphigus vulgaris (50%), however, recognize both dcsmoglein 1 and desmoglein 3 isoforms (Lin *et al*, 1997a; Arteaga *et al*, 2002), and in some cases patients can progress front pemphigus vulgaris to pemphigus foliaccus or vice versa, although the latter is less common (Iwatsuki *et al*, 1991; Kawana *et al*, 1994; Chang *et al*, 1997; Ishii *et al*, 2000). Staphylococcal scalded skin syndrome, also known as pemphi-gus nconatorunr, dermatitis exfoliativa neonatorum, or Ritter's disease, is a generalized skin blistering disease primarily affecting infants and young children. Clinically the disease is characterized by fever, skin tenderness, and erythema with the subsequent loss of large areas of skin due to epidermal separation that occurs in days or even hours of onset. This blistering disease is induced by at least three serotypes of exfoliative toxin (ET) of *Staphylococcus aureus*, namely ETA, ETB, and ETD (Melish and Glasgow, 1970; Melish *et al*, 1972; Hanakawa *et al*, 2002b). Studies have concluded that the various strains of *S. aureus* produce their specific toxins in the skin, whereby they target and cleave the epidermal differentiation-specific desmoglein iso-form, desmoglein 1 (Amagai *et al*, 2000, 2002; Hanakawa *et al*, 2002b). This direct cleavage occurs within the extracellular domain of the protein resulting in dysadhesion between neighboring keratinocytes, thereby destroying the protective barrier of the stratum corneum and allowing the bacteria to enter and flourish within the skin.

As mentioned above the only human desmoglein disease model is the involvement of desmoglein 1 in striate palmoplantar ker-atoderma. No mouse desmoglein 1 models have been successfully engineered to date, however. The sequence homology and expression profile data for desmogleins 5 and 6 make it abundantly clear that any attempt to knock out desmoglein 1 probably results in compensation by desmogleins 5 and 6 within the epidermis. Therefore, methods must be employed to knock out the dcsmoglcin 1, 5, and 6 genes in order to yield a mouse with a superficial epidermal phenotype similar to that of striate palmoplantar kera-toderma. When antibodies that cause human pemphigus foliaccus are injected into mouse skin, however, neither desmoglein 5 nor desmoglein 6 can compensate for the antibody-induced loss of desmoglein 1 as its protein sequence is nearly identical at the predicted antigenic sites, thus resulting in the expected pemphigus foliaceus blistering phenotype. It can also be predicted from homology at the protein level that desmoglein 5 may also be a target for the different forms of exfoliative toxins. These toxins cleave human and mouse desmoglein 1 at a unique site after a glutamic acid residue (i.e., E381 for both human and mouse desmoglein 1) at the carboxy end of extracellular domain 3 (Hanakawa *et al*, 2002b). Indeed, around this cleavage site the amino acid sequence of mouse desmoglein 5 is identical to that of mouse desmoglein 1; however, this sequence is different for mouse desmoglein 6, which lacks the glutamic acid residue. Desmoglein 6 may therefore convey some resistance for the mouse against attack by *S. aureas.*

Of the desmoglein isoforms, desmoglein 3 has been the most widely studied isoform and several mouse models have been constructed. Desmoglein 3 knockout mice demonstrate a loss of keratinocyte cell–cell adhesion that results in a phenotype resembling that of patients with pemphigus vulgaris (Koch *et al*, 1997). Specifically, the oral mucosa demonstrates the full range of lesions seen in individuals with pemphigus vulgaris in addition to suprabasal blisters in traumatized skin. Knockout mice are smaller in size compared with wild-type and heterozygous knockout mice and demonstrate a remarkable decrease in body fat. In addition, after 3–4 wk the mice become bald with a very similar phenotype to balding (bal) mice. Subsequent studies have shown that two naturally occurring autosomal recessive balding (bal and dsg3bal-pas) mice have protein truncating mutations in desmoglein 3 that result in the loss of their cytoplasmic domain (Koch *et al*, 1997; Choi *et al*, 2002). A desmoglein isoform compensation hypothesis has been derived based on several observations (Udey and Stanley, 1999). Neonates of mothers affected with pemphigus foliaceus appear to be protected against the maternal circulating antibodies against desmoglein 1 (Wu *et al*, 2000). The expression of desmoglein 3 in neonatal skin is different to that of human adult skin (Wu *et al*, 2000). In neonates desmoglein 3 is expressed throughout the epidermis whereas, as mentioned above, the expression in adult skin is in the deep epidermis. It therefore seems that the desmoglein 3 expression in the superficial layers of the neonate can compensate for the loss of antibody-induced loss of desmoglein 1. In adult skin, however, desmoglein 3 expression is not present in the superficial layers and hence patients develop blisters throughout the superficial epidermis. The induction of desmoglein 3 expression throughout the superficial epidermis is a very obvious target for gene therapy in patients with pemphigus foliaceus. In another study, it has been demonstrated that expression of desmoglein 1 can compensate for the genetic loss of desmoglcin 3 in keratinocyte adhesion (Hanakawa *et al*, 2002a). Here, the authors directed expression of desmoglein 1 using a keratin 14 promoter in desmoglein 3 knockout mice and studied whether these mice lost their telogen hair prematurely. Although there was some increased hair loss desmoglein 1 could compensate to a certain degree for the lack of desmoglein 3 in the desmosomes (Hanakawa *et al*, 2002a). Very recently, a mouse desmoglein 2 knockout model has been engineered using homologous recombination (Eshkind *et al*, 2002). Both homozygous and a significant number of heterozygous knockout mice died shortly after implantation (Eshkind *et al*, 2002).

Another question that arises with the discovery of three further desmogleins is what their desmocollin binding partners are. A simplistic model based on the expression profiles of these proteins and some binding experiments predicts that desmoglein 1 interacts with desmocollin 1, desmoglein 2 interacts with desmocollin 2, and desmoglein 3 interacts with desmocollin 3. Studies are now required, however, to determine the partners of desmoglcins 4, 5, and 6. Previous studies have shown that desmogleins form weak homophilic interactions (Syed *et al*, 2002) so it can be concluded based on predicted expression patterns that desmoglein 4 probably interacts with desmocollin 3 within the basal and immediate suprabasal cells, and that desmoglcins 5 and 6 interact with desmocollin 1 in the superficial layers.

Apart from their adhesive capacity, do the dcsmosomal cadher-ins perform any additional roles? In a recent study to determine what might happen when a desmosomal cadherin is expressed out of its correct cell layer Ishii *et al* (2001) discovered by expression of ectopic desmosomal

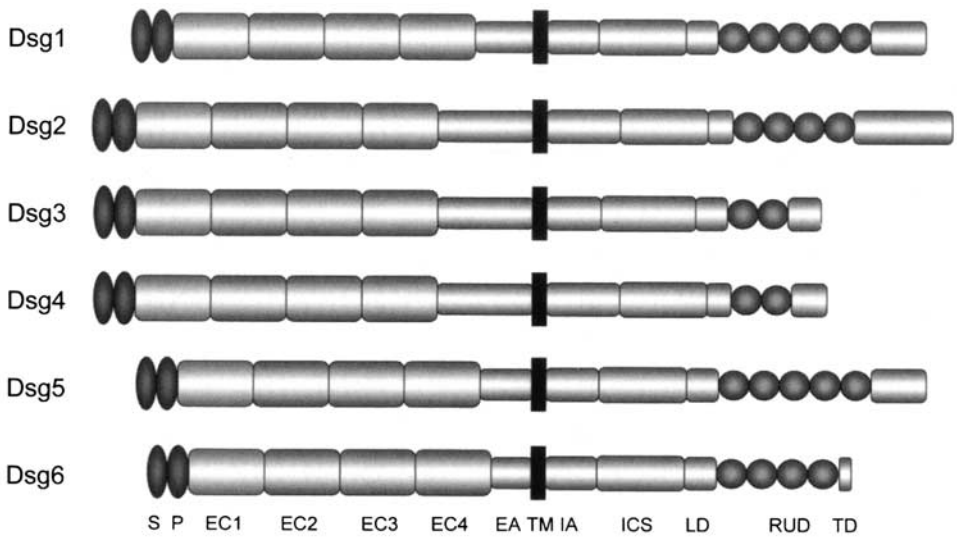

Figure 3. Schematic representation of the six mouse encoded proteins. Each isoform comprises a signal (S) and preprotein (p) domain followed by five extracellular domains (EC1–4, EA), a transmembrane domain (TM), an intracellular anchor domain (IA), an intracellular cadherin-typical segment domain (ICS), a linker domain (LD), a repeat unit domain (RUD) containing various repeats, and a terminal domain (TD).

Table X. Molecular weight and isoelectric points for precursor and cleaved mature mouse desmoglein isoforms

	Precursor proteins			Cleaved proteins		
Isoform	Number of amino acids	Molecular weight (Da)	Isoelectric point (pi)	Number of amino acids	Molecular weight (Da)	Isoelectric point (pi)
Dsg1	1057	114,597	4.78	1008	108,771	4.64
Dsg2	1122	122,385	5.16	1068	116,324	5.01
Dsg3	993	107,902	4.82	944	102,448	4.67
Dsg4	1041	114,449	4.58	992	108,598	4.55
Dsg5	1060	114,454	4.72	1011	108,628	4.59
Dsg6	911	100,516	4.84	862	94,690	4.68

Mouse desmoglein 1, Mahoney *et al* (2002); mouse desmoglein 2, Mahoney *et al* (2002); mouse desmoglein 3, Ishikawa *et al* (2000).

cadherins that epithelial cells can tolerate large ranges in the amounts of endogenous dcsmosomal cadherins, but that low levels expressed out of their normal cellular context (i.e., desmoglein 1) can disrupt desmosomes. They concluded that there must be a specific program of differentiation in which the spatial and temporal expression pattern of the differentiation-specific dcsmosomal cadherins (i.e., desmoglein 1 and desmocollin 1) is coordinated with that of their binding partners, which facilitate their incorporation into desmosomes (Ishii *et al*, 2001). In another study to determine whether the distribution of the different dcsmoglcin isoforms affected the structure or function of the epidermis, transgenic mice were bred with desmoglein 3 driven by the involucrin promoter to express the transgcnc in the superficial epidermis (Elias *et al*, 2001). These mice demonstrated an abnormal stratum corneum that was more like that of mucous membranes and showed an abnormal skin barrier function characterized by increased transcpidcrmal water loss (Elias *et al*, 2001). It can therefore be concluded from all of these studies that, in addition to the obvious interchangeable adhesive role of the desmoglein isoforms, their relative ratios within any layer in the skin dictate the precise structure and function of those keratinocytes. Interestingly, RT-PCR demonstrated that desmoglein 2 was widely expressed in desmosome and nondesmosome containing tissues (**Table VI**). Could this be due to the oversensitive nature of the RT-PCR used or docs desmoglein 2 play an important role in nondesmosomal tissues? Recent studies have shown that desmoglein 2 is essential for the inner cell mass of blastocysts and the embryonal stem cell population derived from them (Eshkind *et al*, 2002). These authors have shown that desmoglein 2 is located in desmoplakin-negative wild-type embryonal stem cells in nondesmosomal junctions and demonstrate that it is required for normal embryonal stem cell proliferation (Eshkind *et al*, 2002). Therefore, it can be concluded that desmoglein 2 may play a much larger role in mouse development than just as an attachment molecule.

In summary, we have demonstrated that there are six mouse desmoglein genes. Studies are now required to determine the expression profile of these isoforms at the protein level, to determine their binding partners, and to use recombinant mouse technology in order to determine the essential roles of the proteins. These results may have some impact on the desmoglein compensation hypothesis in mouse and in addition may have some value in human gene therapy in diseases such as pemphigus.

ELECTRONIC DATABASE INFORMATION

URLs for data in this article are as follows: Online Mendelian Inheritance in Man (OMIM), http://www3.ncbi.nhn.nih.gov/Omim/
Human Genome Working Draft, http://genome.ucsc.edu/
The Protein Machine, http://www2.ebi.ac.uk/ Biology WorkBcnch 3.2, http://biowb.sdsc.edu/
Expert protein analysis system, http://us.expasy.org/

ACKNOWLEDGMENTS

This work was funded by Action Research (salary to NW). We also wish to thank the Royal Devon and Exeter Hospital.

REFERENCES

Adams MJ, Reichel MB, King IA *et al* (1998) Characterization of the regulatory regions in the human desmoglcin genes encoding the pemphigus foliaceous and pemphigus vulgaris antigens. *Biochem J* 329:165–74

Amagai M, Klaus-Kovtun V, Stanley JR (1991) Autoantibodies against a novel epithelial cadherin in pemphigus vulgaris, a disease of cell adhesion. *Cell* 67:869–77

Amagai M, Matsuyoshi N, Wang ZH, Andl C, Stanley JR (2000) Toxin in bullous impetigo and staphylococcal scalded-skin syndrome targets dcsmoglein 1. *Nat Med* 6:1275–7

Amagai M, Yamaguchi T, Hanakawa Y, Nishifuji K, Sugai M, Stanley JR (2002) Staphylococcal exfoliative toxin B specifically cleaves desmoglein 1. *J Invest Dermatol* 118:845–50

Andl CD, Stanley JR (2001) Central role of the plakoglobiu-binding domain for desmoglein 3 incorporation into desmosomes. *J Invest Dermatol* 117:1068–74

Armstrong DK, McKenna KE, Purkis PE, Green KJ, Eady RA, Leigh IM, Hughes AE (1999) Haploinsufficiency of desmoplakin causes a striate subtype of palmoplantar keratoderma. *Hum Mol Genet* 8:143–8

Arnemann J, Sullivan KH, Magee AI, King IA, Buxton RS (1993) Stratification-related expression of isoforms of the desmosomal cadherins in human epidermis. *J Cell Set* 104:741–50

Arteaga LA, Prisayanh PS, Warren SJ, Liu Z, Diaz LA, Lin MS (2002) A subset of pemphigus foliaceus patients exhibits pathogenic autoantibodies against both desmoglein-1 and desmoglcin-3. *J Invest Dermatol* 118:806–11

Bannon LJ, Cabrera BL, Stack MS, Green KJ (2001) Isoform-specific differences in the size of desmosomal cadherin/catenin complexes. *J Invest Dermatol* 117:1302–6

Buxton RS, Cowin P, Franke WW *et al* (1993) Nomenclature of the desmosomal cadherins. *J Cell Biol* 121:481–3

Chang SN, Kim SC, Lee IJ, Seo SJ, Hong CK, Park WH (1997) Transition from pemphigus vulgaris to pemphigus foliaceus. *BrJ Dermatol* 137:303–5

Chidgey MA, Yue KK, Gould S, Byrne C, Garrod DR (1997) Changing pattern of desmo collin 3 expression accompanies epidermal organisation during skin development. *Dev Dyn* 210:315–27

Chitaev NA, Troyanovsky SM (1997) Direct Ca^{2+}-dependent heterophilic interaction between desmosomal cadherins, desmoglcin and desmocollin contributes to cell-cell adhesion. *J Cell Biol* 138:193–201

Choi Y, Pulkkincn L, Simpson A, Montagutelli X, Sundbcrg J, Uitto J, Maboncy MG (2002) Loss of cell adhesion in Dsg3bal-Pas mice with homozygous deletion mutation (2079dell4) in the desmoglcin 3 gene. *J Invest Dermatol* 119:226

Denning MF, Guy SG, Ellcrbroek SM, Norvell SM, Kowalczyk AP, Green KJ (1998) The expression of dcsmoglein isoforms in cultured human kcratinocytcs is regulated by calcium, serum, and protein kinase C. *Exp Cell Res* 239:50–9

Diaz LA, Sampaio SA, Rivitti EA *et al* (1989) Endemic pemphigus foliaceus (fogo selva-gem). I. Clinical features and imrnunopathology. *J Am Acad Dermatol* 20:657–69

Elias PM, Matsuyoshi N, Wu H, Lin C, Wang ZH, Brown BE, Stanley JR (2001) Desmoglein isotorm distribution affects stratum corneum structure and function. *J Cell Biol* 153:243–9

El Saved F, Bazcx J (1993) Kcratodcrmie pahnoplantaire striee. *Ann Dermatol Venereal* 120:894–5

Emery DJ, Diaz LA, Fairley JA, Lopez A, Taylor AF, Giudice GJ (1995) Pemphigus foliaceus and pemphigus vulgaris autoantibodies react with the extracellular domain of desmoglein-1. *J Invest Dermatol* 104:323–8

Eshkind L, Tian Q, Schmidt A, Franke WW, Windoffer R, Leube RE (2002) Loss of desmoglcin 2 suggests essential functions for early embryonic development and proliferation of embryonal stem cells. *Eur J Cell Biol* 81:592–8

Frank J, Cscrhalmi-Fricdman PB, Ahmad W, Panteleycv AA, Aita VM, Christiano AM (2001) Characterization *of* the desmosomal cadherin gene family: Genomic organization of two desmoglein genes on human chromosome l8q12. *Exp Dermatol* 10:90–4

Griffiths W, Judge M, Leigh I (1998) Disorders of keratinization. In: Rook A, Wilkinson D, Ebling F (eds). *Textbook of Dermatology*. Oxford: Blackwell pp 1483–588

Hanakawa Y, Matsuyoshi N, Stanley JR (2002a) Expression of desmoglcin 1 compensates for genetic loss of desmoglcin 3 in kcratinocyte adhesion. *J Invest Dermatol* 119:27–31

Hanakawa Y, Schechter NM, Lin C *et al* (2002b) Molecular mechanisms of blister formation in bullous impetigo and staphylococcal scalded skin syndrome. *J Clin Invest* 110:53–60

Hunt DM, Rickman L, Whittock NV *et al* (2001) Spectrum of dominant mutations in the desmosomal cadherin desmoglcin 1, causing the skin disease striate palmoplan-tar keratoderma. *Eur J Hum Genet* 9:197–203

Ishii K, Amagai M, Ohata Y, Shimizu H, Hashimoto T, Ohya K, Nishikawa T (2000) Development *of* pemphigus vulgaris in a patient with pemphigus foliaceus: Antidesmoglcin antibody profile shift confirmed by enzyme-linked immunosorbent assay. *J Am Acad Dermatol* 42:859–61

Ishii K, Norvell SM, Bannon LJ, Amargo EV, Pascoe LT, Green KJ (2001) Assembly of desmosomal cadherins into desmosomes is isoform dependent. *J Invest Dermatol* 117:26–35

Ishikawa H, Li K, Tamai K, Sawamura D, Uitto J (2000) Cloning of the mouse dcsmoglein 3 gene (Dsg3): Interspecies conservation within the cadherin supcrfamily. *Exp Dermatol* 9:229–39

Iwatsuki K, Takigawa M, Hashimoto T, Nishikawa T, Yamada M (1991) Can pemphigus vulgaris become pemphigus foliaceus? *J Am Acad Dermatol* 25:797–800

Kapprell HP, Owaribe K, Franke WW (1988) Identification of a basic protein of Mr 75,000 as an accessory desmosomal plaque protein in stratified and complex epithclia. *J Cell Biol* 106:1679–91

Kawana S, Hashimoto T, Nishikawa T, Nishiyama S (1994) Changes in clinical features, histologic findings, and antigen profiles with development of pemphigus foliaceus from pemphigus vulgaris. *Arch Dermatol* 130: 1534–1538

King IA, Arnemann J, Spurr NK, Buxton RS (1993) Cloning of the cDNA (DSC1) coding for human type 1 desmocollin and its assignment to chromosome 18. *Genomics* 18:185–94

King IA, Sullivan KH, Bennett R Jr, Buxton RS (1995) The desmocollins of human foreskin epidermis: Identification and chromosomal assignment of a third gene and expression patterns of the three isoforms. *J Invest Dermatol* 105:314–21

King IA, O'Brien TJ, Buxton RS (1996) Expression of the 'skin-type' desmosomal cadherin DSC1 is closely linked to the keratinization of epithelial tissues during mouse development. *J Invest Dermatol* 107:531–8

Kljuic A, Gilead L, Martinez-Mir A, Frank J, Christiano AM, Zlotogorski A (2003) A nonsense mutation in the desmoglein 1 gene underlies striate keratoderma. *Exp Dermatol*

Koch PJ, Walsh MJ, Schmclz M, Goldschmidt MD, Zimbelmann R, Franke WW (1990) Identification of dcsmoglein, a constitutive desmosomal glycoprotein, as a member of the cadherin family of cell adhesion molecules. *Eur J Cell Biol* 53:1–12

Koch PJ, Goldschmidt MD, Walsh MJ, Zimbelmann R, Franke WW (1991) Complete amino acid sequence of the epidermal desmoglcin precursor polypeptide and identification of a second type of desmoglein gene. *Eur J Cell Biol* 55:200–8

Koch PJ, Goldschmidt MD, Zimbelmann R, Troyanovsky R, Franke WW (1992) Complexity and expression patterns of the desmosomal cadherins. *Proc Natl Acad Sci USA* 89:353–7

Koch PJ, Mahoney MG, Ishikawa H *et al* (1997) Targeted disruption of the pemphigus vulgaris antigen (desmoglcin 3) gene in mice causes loss of keratinocyte cell adhesion with a phenotype similar to pemphigus vulgaris. *J Cell Biol* 137:1091–102

Kowalczyk AP, Borgwardt JE, Green KJ (1996) Analysis of desmosomal cadberin-adhesive function and stoichiometry of desmosomal cadherin–plakoglobin complexes. *J Invest Dermatol* 107:293–300

Labib RS, Rock B, Martins CR, Diaz LA (1990) Pemphigus foliaceus antigen: Characterization of an immunoreactivc tryptic fragment from BALB/c mouse epidermis recognized by all patients' sera and major autoantibody subclasses. *Clin Immunol Immunopathol* 57:317–29

Lever WF (1965) Pemphigus vulgaris. In: Lever WF (ed). *Pemphigus and Pemphigoid*. 1st edn. Springfield, IL: Charles C. Thomas

Lin MS, Swartz SJ, Lopez A, Ding X, Fairley JA, Diaz LA (1997a) T lymphocytes from a subset of patients with pemphigus vulgaris respond to both dcsmoglein-3 and desmoglein-1. *J Invest Dermatol* 109:734–7

Lin MS, Swartz SJ, Lopez A *et al* (1997b) Development and characterization of dcsmoglein-3-spccificTcells from patients with pemphigus vulgaris. *J Clin Invest* 99:31–40

Lin MS, Fu CL, Aoki V *et al* (2000) Desmoglein-1-specific T lymphocytes from patients with endemic pemphigus foliaceus (fogo selvagcm). *J Clin Invest* 105:207–13

Mahoney MG, Simpson A, Aho S, Uitto J, Pulkkincn L (2002) Interspecies conservation and differential expression of mouse desmoglein gene family. *Exp Dermatol* 11:115–25

Marcozzi C, Burdctt ID, Buxton RS, Magee AI (1998) Cocxpression of both types of desmosomal cadherin and plakoglobin confers strong intercellular adhesion. *J Cell Sci* 111:495–509

Mathur M, Goodwin L, Cowin P (1994) Interactions of the cytoplasmic domain of the desmosomal cadherin Dsgl with plakoglobin. *J Biol Chem* 269:14075–80

Mclish ME, Glasgow LA (1970) The staphylococcal scalded-skin syndrome. *N Engl J Med* 282:1114–9

Mclish ME, Glasgow LA, Turner MD (1972) The staphylococcal scalded-skin syndrome: Isolation and partial characterization of the exfoliative toxin. *J Infect Dis* 125:129–40

Nilles LA, Parry DA, Powers EE, Angst BD, Wagner RM, Green KJ (1991) Structural analysis and expression of human desmoglein: A cadherin-like component of the desmosome. *J Cell Sci* 99:809–21

Nollet F, Kools P, van Roy F (2000) Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. *J Mol Biol* 299:551–72

North AJ, Chidgcy MA, Clarke JP, Bardsley WG, Garrod DR (1996) Distinct desmocollin isoforms occur in the same desmosomes and show reciprocally graded distributions in bovine nasal epidermis. *Proc Natl Acad Sci USA* 93:7701–5

North AJ, Bardsley WG, Hyam J *et al* (1999) Molecular map of the desmosomal plaque. *J Celt Sci* 112:4325–36

Nuber UA, Schafer S, Schmidt A, Koch PJ, Frankc WW (1995) The widespread human desmocollin Dsc2 and tissue-specific patterns of synthesis of various desmocollin subtypes. *EurJ Cell Biol* 66:69–74

Olaguc-Alcala M, Giudice GJ, Diaz LA (1994) Pemphigus foliaceus sera recognize an N-terminal fragment of bovine desmoglein 1. *J Invest Dermatol* 102:882–5

Ourslcr JR, Labib RS, Anss-Abdo L, Burke T, O'Keefe EJ, Anhalt GJ (1992) Human autoantibodies against desmoplakins in paraneoplastic pemphigus. *J Clin Invest* 89:1775–82

Ozawa M, Terada H, Pedraza C (1995) The fourth armadillo repeat of plakoglobin (γ-catenin) is required for its high affinity binding to the cytoplasmic domains of E-cadherin and desmosomal cadhcrin Dsg2, and the tumor suppressor APC protein. *JBiochem (Tokyo)* 118:1077–82

Puttagunta S, Mathur M, Cowin P (1994) Structure of DSG1, the bovine desmosomal cad-herin gene encoding the pemphigus foliaceus antigen. Evidence of polymorphism. *J Biol Chem* 269:1949–55

Rickman L, Simrak D, Stevens HP *et al* (1999) N-terminal deletion in a desmosomal cad-herin causes the autosomal dominant skin disease striate palmoplantar kerato-derma. *Hum Mol Genet* 8:971–6

Schafer S, Koch PJ, Frankc WW (1994) Identification of the ubiquitous human desmoglein, Dsg2, and the expression catalogue of the desmoglein subfamily of desmosomal cadherins. *Exp Cell Res* 211:391–9

Schafer S, Stumpp S, Frankc WW (1996) Immunological identification and characterization of the desmosomal cadhcrin Dsg2 in coupled and uncoupled epithelial cells and in human tissues. *Differentiation* 60:99–108

Schmclz M, Duden R, Cowin P, Frankc WW (1986) A constitutive transmembrane glycoprotein of Mr 165,000 (desmoglein) in epidermal and non-epidermal desmosomes. I. Biochemical identification of the polypeptide. *Eur J Celt Biol* 42:177–83

Schmidt A, Heid HW, Schafer S, Nuber UA, Zimbelmann R, Franke WW (1994) Desmosomes and cytoskeletal architecture in epithelial differentiation: Cell type-specific plaque components and intermediate filament anchorage. *Eur] Cell Biol* 65:229–45

Siemens H (1929) Keratosis palmo-plantaris striata. *Arch Dermatol Syphiloi* 157:392–408

Silos SA, Tamai K, Li K, Kivirikko S, Kouba D, Christiano AM, Uitto J (1996) Cloning of the gene for human pemphigus vulgaris antigen (desmoglein 3), a desmosonul cadherin. Characterization of the promoter region and identification ofaker-atinocyte-specific ci5-element. *J Biol Chem* 271:17504–11

Simrak D, Cowley CM, Buxton RS, Arnemann J (1995) Tandem arrangement of the closely linked desmoglein genes on human chromosome 18. *Genomics* 25:591–4

Stanley JR, Koulu L, Thivolct C (1984) Distinction between epidermal antigens binding pemphigus vulgaris and pemphigus foliaceus autoantibodies. *J Clin Invca* 74:313–20

Syed SE, Trinnaman B, Martin S, Major S, Hutehinson J, Magee AI (2002) Molecular interactions between desmosomal cadherins. *BiochemJ* 362:317–27

Theis DG, Koch PJ, Franke WW (1993) Differential synthesis of type 1 and type 2 desmocollin mRNAs in human stratified epithelia. *Int J Dev Biol* 37:101–10

Troyanovsky RB, Chitaev NA, Troyanovsky SM (1996) Cadherin binding sites of plakoglobin: Localization, specificity and role in targeting to adhering junctions. *J Cell Sci* 109:3069–78

Troyanovsky SM, Eshkind LG, Troyanovsky RB, Leube RE, Franke WW (1993) Contributions of cytoplasmic domains of desmosomal cadherins to desmo-some assembly and intermediate filament anchorage. *Cell* 72:561–74

Troyanovsky SM, Troyanovsky RB, Eshkind LG, Krutovskikh VA, Leubc RE, Franke WW (1994a) Identification of the plakoglobin-binding domain in desmoglein and its role in plaque assembly and intermediate filament anchorage. *J CellE* 127:151–60

Troyanovsky SM, Troyanovsky RB, Eshkind LG, Leubc RE, Franke WW (1994b) Identification of amino acid sequence motifs in desmocollin, a desmosomal glycoprotein, that are required for plakoglobin binding and plaque formation. *Proc Natl Acad Sci USA* 91:10790–4

Tselepis C, Chidgey M, North A, Garrod D (1998) Desmosomal adhesion inhibits invasive behavior. *Proc Natl Acad Sci USA* 95:8064–9

Udey MC, Stanley JR (1999) Pemphigus – diseases of antidesmosomal auto-immunity. *Jam* 282:572–6

Wahl JK, Sacco PA, McGranahan-Sadler TM, Sauppe LM, Wheelock MJ, Johnson KR (1996) Plakoglobin domains that define its association with the desmosomal cadherins and the classical cadherins: Identification of unique and shared domains. *J Cell Sci* 109:1143–54

Wheeler GN, Parker AE, Thomas CL *et al* (1991) Desmosomal glycoprotein DGI, a component of intercellular desmosome junctions, is related to the cadhcrin family of cell adhesion molecules. *Proc Natl Acad Sci USA* 88:4796–800

Whittock NY, Bower C (2003) Genetic evidence for a novel human desmosomal cadhcrin. desmoglein 4. *J Invest Dermatol* 120:523–30

Witeher LL, Collins R, Puttagunta S, Mechanic SE, Munson M, Gumbincr B, Cowin P (1996) Desmosomal cadherin binding domains of plakoglobin. *J Biol Chem* 271:10904–9

Wu H, Wang ZH, Yan A *et al* (2000) Protection against pemphigus foliaceus by desmoglein 3 in neonates. *N Engl J Med* 343:31–5

Yue KK, Holton JL, Clarke JP, Hyam JL, Hashimoto T, Chidgey MA, Garrod DR (1995) Characterisation of a desmocollin isoform (bovine DSC3) exclusively expressed in lower layers of stratified epithelia. *J Cell Sci* 108:2163–73